DOCUMENT RESUME

ED 253 843 CS 007 912

AUTHOR Nolen, Patricia A.; Lam, Tony C. M.
TITLE A Comparison of IRI and Durrell Analysis of Reading Difficulty Reading Levels in Clinical Assessment.
PUB DATE [81]
NOTE 13p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Clinical Diagnosis; *Comparative Analysis; Elementary Education; *Informal Reading Inventories; *Reading Diagnosis; Reading Difficulties; Reading Rate; *Reading Research; *Reading Tests; *Scoring; Test Reliability; Test Validity
IDENTIFIERS *Durrell Analysis of Reading Difficulty

ABSTRACT

A comparison was made of Durrell Analysis of Reading Difficulty and informal reading inventory (IRI) independent and instructional level designations for 15 children, aged 9 to 11 years, who had been referred to a diagnostic clinic for reading assessment. The children's reading performance was first scored according to procedures outlined in the Durrell Analysis manual, and a second scoring was made according to recommendations for administering informal reading inventories given by M. S. Johnson and R. A. Kress. Results suggested that the procedures yield significantly different overall grade level designations. Further analysis of the significant reading level by scoring procedure interaction showed that (1) the average independent level established by the Durrell Analysis procedure was significantly higher than that obtained by the IRI procedure, and (2) the instructional level mean was significantly higher than the independent level mean only when the IRI procedure was used. (FL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

ED253843

A Comparison of IRI and Durrell Analysis of Reading Difficulty

Reading Levels in Clinical Assessment

Patricia A. Nolen

Tony C. M. Lam

University of Washington

College of Education

Running Head: A Comparison of IRI and Durrell Analysis of Reading

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Patricia Nolen

Tony C. M. Lam

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

CS 007 912

Abstract

A 2 X 2 ANOVA with repeated measures comparison was made of 15 clinical independent and instructional reading level assignments of Informal Reading Inventory (IRI) and the revised Durrell Analysis of Reading Difficulty which includes reading rate as a scoring criterion. Results suggested that the procedures yield significantly different overall grade level designations. Further analysis of the significant reading level by scoring procedure interaction showed that (a) the average independent level established by the Analysis procedure was significantly higher than that obtained by the IRI procedure, and (b) the instructional level mean was significantly higher than the independent level mean only when the IRI procedure was employed. Findings also suggested that overall Analysis derived levels admit usual interpretation for parents and teachers.

A Comparison of IRI and Durrell Analysis of Reading Difficulty Reading Levels in Clinical Assessment

Various understandings of reading and methods of testing may yield different diagnostic assessments, especially where they modify what can be measured. Traditionally, clinical reading includes a graded series of paragraphs for oral reading by a child with little preparation as far as the topic or word ideas within. As the child reads, an examiner makes an accounting of departures from the printed text according to criteria that frequently vary with an examiner's understanding of the process, and/or the particular error classification scheme used. (Jongsma & Jongsma, 1981). Although reader's word changes relative to the text are not difficult to note, word recognition error scores are not derived by simple addition. A misreading of the printed word <u>were</u> as <u>was</u> may or may not be a scorable error, depending not only on the reader's characteristic speech patterns, but also on whether or not speech patterned substitutions figure in an examiner's understanding (Goodman, 1977).

Further complicating precision in error counts are classification schemes that assign different weights to words read that do not literally duplicate words in print. Should the child in the example return to the original <u>were</u>/<u>was</u> misreading and correct the first pronunciation, one classification scheme allots the correction full credit (Rupley & Blair, 1980); another, half credit (Durrell & Catterson, 1980); and a third may count the repetition itself as an error (Ekwall, 1976). No less judgmental than scoring word accuracy are procedures for scoring tests of comprehension. Where children answer a test question with a key word they have

misread, their answer is correct from one point of view as they comprehended and remembered what they read, but incorrect from another as what they read and remembered was wrong.

Differences in understanding reading, along with differences in classifying errors, might seem trivial were it not common practice to determine graded levels for children's instructional and independent reading, based on word identification and recall of paragraph content. Informal Inventory guidelines traditionally require 99% word recognition and 95% comprehension adequacy in establishing an independent level, and 95% word recognition and 75% comprehension adequacy for instructional purposes (Johnson & Kress, 1965). While there is some controversy over the stringency of the criteria, the validity of level assignments, and the extent to which criteria may be adjusted depending on type of error and reading age (Jongsma & Jongsma, 1981), clinical diagnosis usually adheres to conservative guidelines like those reviewed by Johnson and Kress (1965). Thus, the number of allowable errors in traditional inventories is based on percentages, although a recent review finds no single standard for evaluating misreadings as scorable (Jongsma & Jongsma, 1981).

The rate of reading a graded selection does not often figure prominently in determining reading levels. Although assessments usually include timing or words-per-minute for normative (Spache, 1972) or oral/silent reading comparisons, references to reading rate are just as likely to be simple notations of a readers' fluency and phrasing. Rate, of course,

varies with purpose, and children's purposes in a test situation vary from determination to proceed slowly and avoid error, to more normal pacing, to reading as fast as possible either to impress the listener or end the test. Thus, for most children a general evaluation of reading rate coupled with observations on their task approach would seem all that were needed.

On the other hand, clinical diagnosis deals with severely impaired readers, for whom slow processing of language codes now seems characteristic (Vellutino, 1979). While stopwatch timing probably is not sufficiently sensitive for revealing any abstract processing lags of such readers, it nevertheless might prove useful in assigning levels, if only as a measure of reader effort. Where adequate recall and word identificiation take an inordinate amount of time or labor, a child is not likely to profit from material of similar difficulty for either spontaneous reading or study. Including rate as a supplementary criterion for establishing appropriate levels for remedial planning would seem to be especially pertinent for children referred to clinics.

Moreover, unlike error classification schemes, reading rate standards offer reduced opportunity for interpretive adjustments. If 50 seconds is the norm for reading a test paragraph, then 65 seconds for reading of the same paragraph is clearly outside the allowable time, whatever word or comprehension errors are made or however they are scored. Thus, reading rate criteria as part of error classification schemes suggest more rigorous guidelines for establishing reading levels than do schemes without rate standards.

A test frequently included in clinic batteries is the Durrell Analysis of Reading Difficulty. Newly revised and standardized (Durrell & Catterson, 1980), the test now offers normative data with rate norms for each graded paragraph selection. Primarily an instrument for evaluation and study of an individual reader's strengths and weaknesses, Analysis assessment procedures are not directly comparable to traditional reading inventories in establishing instructional and independent reading levels. Whereas inventories advocate basing a child's reading levels on the accuracy of oral reading and recall while taking into account any discrepancies in silent reading, the Analysis uses time as its major criterion for level assignments, qualified by word recognition and/or comprehension adequacy where appropriate.

For example, an Analysis instructional reading level is obtained from that graded paragraph a child reads aloud within the time norms, with fewer than seven word recognition errors, and a comprehension score based on no more than three errors out of 6-9 test questions (50-66%). The independent level is established by silent reading of graded paragraphs completed within a designated time and for which recall of literal text portions called "Memories" meets a given count. As in most assessment procedures, certain adjustments are advocated where reading performances are not consistent with passages of increasing complexity. Moreover, unlike a single passage equivalent to a sampling of text at one grade level as is characteristic of most tests, Analysis passages translate to different grade levels depending on reading rate. Thus, if the silent reading with

adequate recall of one Analysis passage takes up to 58 seconds, a child's performance is placed within 2nd grade, whereas shorter reading times may designate 3rd or 4th grade level achievement for the same selection. While a number of passages are administered to confirm performance levels, Analysis methods of scoring are not devoid of examiner judgment, despite provision of exact reading time standards for each selection. Nevertheless, Analysis procedures are intuitively appealing for clinical use, given the changing nature of our understanding of reading, differences in error classification schemes, and a current emphasis on time as the appropriate measure for assessing underlying mental processes having to do with language (Sternberg, 1980).

Despite the appeal of time-structured assessments, it is possible that procedural differences in scoring may yield different instructional and independent reading levels. As clinic evaluations include recommendations of graded materials for children's school use, some basis for evaluating the outcomes of the two scoring methods is needed.

A comparison was made of Analysis and Inventory derived independent and instructional level designations for 15 children ages 9 to 11, referred to the University of Washington Psycho-Education Clinic for reading assessment. The children had been categorized as severely reading disabled, functioning 1½ to 2 years below expectancy according to estimates based on years in school and measured intelligence. Their WISC full scales IQs ranged from 87 to 126 with a mean of 107.5, and standard deviation of 1.96. Excluded from the sample were those children whose neurological

functioning and/or emotional adjustment were considered a major factor in their learning disability. The children's reading performances were scored first according to procedures outlined in the Analysis manual. A second scoring of their test protocols was made according to recommendations for administering traditional informal reading inventories given in Johnson and Kress (1965).

The dual scoring was not accomplished without some compromise. Inventory techniques take into account error percentages in comprehension and word recognition, while Analysis methods supply numerical error allowances. Further, as Analysis paragraphs may have no single within grade interpretation, their H (high), M (medium), and L (low) grade scores were translated as .7, .5, and .3 respectively for east of computation. Where percentage of error counts differed from Analysis allowances, Inventory grade levels were assigned by stepped adjustments. Thus, if a middle 2nd grade (2.5) reading performance by Analysis scoring did not meet Inventory word recognition or comprehension standards, the Inventory grade level was designated as low 2nd grade (2.3) or high first grade (1.7), according to whether one or both of the traditional percentage measures differed.

The means ($\bar{X}$) and standard deviations (S) of graded levels for the subjects' independent and instructional readings derived by the Analysis and Inventory scoring methods were computed and summarized in Table 1.

Insert Table 1 about here

Product moment correlations among the four measures of reading levels were calculated and the range was from 0.9053 to 0.9527, all significant at the 0.01 level. In order to determine the statistical significance of the differences in means, a 2 X 2 ANOVA with repeated measures design, with reading level and scoring procedure as the two factors, was used to analyze the data. Results indicated that there was no significant difference between the reading levels ($F_{1,14} = 1.3$, $p > .05$), but a significant difference between the scoring procedures ($F_{1,14} = 8.4$, $p < .05$) and a significant interaction between reading levels and scoring procedures ($F_{1,14} = 9.4$, $p < .01$). In order to test the interaction further, paired t-tests were calculated to assess the simple main effects. When comparing the independent reading levels derived by the two scoring methods, the Analysis procedure had a significantly higher mean graded level than the Inventory procedure ($t_{14} = 3.80$, $< .01$). No such significant difference in means was found between the two scoring methods for the instructional levels ($t_{14} = 0.48$, $p > .05$). With the Inventory scoring procedure, the average instructional reading level derived was significantly higher than the average independent reading level ($t_{14} = 2.39$, $p < .05$) as would be expected. However, the means of the two reading levels were not significantly different from each other ($t_{14} = 1.01$, $p > .05$) when the Analysis scoring procedure was employed.

In summary, the correlational data analysis indicated that a person's reading performance will be rated similarly in relation to other children's performances by both scoring procedures in both reading levels. However, the findings from comparing means suggested that a child's absolute graded level is somewhat different depending on whether the Analysis or Inventory scoring method is used. If the Analysis is employed, a child's independent reading level will tend to be rated higher, and the difference between the reading levels will tend to be smaller. On the other hand, if the Inventory procedure is used, a child's independent reading level will tend to be rated lower, and the difference between the reading levels will tend to be greater.

Although results must be considered preliminary as Inventory levels were not completely independent of Analysis levels, they underscore the importance of appraising the contribution of assessment techniques in differential reading diagnosis. Results suggest first, that Analysis outcomes need no more (or less) interpretation of their approximation to children's grade level function than do Inventory outcomes. Second, they suggest that time as a criterion yields very similar graded levels for instruction and independent reading. Third, an Analysis determination of an independent level may not be as conservative as an Inventory designation. In planning for children whose reading difficulties have brought them to clinic attention, this might be taken into account. However, firm recommendations await further study as ad hoc grade level inventory assignments were necessitated by characteristics of the Analysis test passages on which they were computed.

References

Durrell, D. D., & Catterson, J. H. Durrell analysis of reading difficulty (2nd ed.). New York: The Psychological Corporation, 1980.

Ekwall, E. E. Diagnosis and remediation of the disabled reader. Boston: Allyn and Bacon, 1976.

Goodman, K. S. Miscues: Windows on the reading process. Miscue analysis. Urbana, Illinois: ERIC Clearinghuse on Reading and Communicating Skills, National Council of Teachers of English, 1976, pp. 3-14.

Johnson, M. S, & Kress, R. A. Informal reading inventories. Newark, Delaware: International Reading Association, 1965.

Jongsma, K. S., & Jongsma, E. A. Test review: Commercial informal reading inventories. Reading Teacher, 1981, 34(6), 697-705.

Rupley, W. H., & Blair, T. R. Reading diagnosis and remediation. Chicago: Rand McNally, 1979.

Spache, G. D. Diagnostic reading scales. Examiner's Manual. Revised Edition. Monterey, California: CTB/McGraw Hill, 1972.

Sternberg, R. J. The nature of mental abilities. American Psychologist, 1979, 34, 214-230.

Vellutino, F. R. Dyslexia: Theory and research. Cambridge, Mass.: The MIT Press, 1979.

Table 1

	Analysis		Inventory	
	Independent	Instructional	Independent	Instructional
$\bar{X}$	3.067	2.953	2.547	2.900
S	1.428	1.398	1.442	1.399